OPEN ACCESS

Data Paper

The InBIO Barcoding Initiative Database: contribution to the knowledge of DNA barcodes of the vascular plants of north-eastern Portugal

João Queirós[‡,§,|,¶], Rodrigo Silva[|], Catarina J. Pinho[‡,§,|], Hélia M. Vale-Gonçalves[#,¤], Ricardo Pita[«], Paulo C. Alves[‡,§,|,»], Pedro Beja[‡,^], Joana Paupério[‡,ˇ], Miguel Porto[‡,¦,^,»]

‡ CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Campus de Vairão, Universidade do Porto, 4485-661 Vairão, Vila do Conde, Portugal

§ BIOPOLIS, Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Vila do Conde, Portugal

| Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, Rua do Campo Alegre s/n, 4169-007 Porto, Portugal

¶ Mértola Biological Station, 7750-329, Mértola, Portugal

CITAB-Inov4Agro, Centro de Investigação e Tecnologias Agroambientais e Biológicas, Universidade de Trás-os-Montes e Alto Douro (UTAD), 5000-801 Vila Real, Portugal

¤ LEFT, Laboratório de Ecologia Fluvial e Terrestre, UTAD, Vila Real, Portugal

« Instituto Mediterrâneo para a Agricultura, Ambiente e Desenvolvimento (MED), Universidade de Évora, Pólo da Mitra, 7002-554 Évora, Portugal

» Mértola Biological Station, 7750-329 Mértola, Portugal

^ BIOPOLIS Program in Genomics, Biodiversity and Land Planning, CIBIO, Campus de Vairão, 4485-661 Vairão, Vila do Conde, Portugal

ˇ Current address: European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, CB10 1SD, United Kingdom

¦ CIBIO, Centro de Investigação em Biodiversidade e Recursos Genéticos, InBIO Laboratório Associado, Instituto Superior de Agronomia, Universidade de Lisboa, 1349-017 Lisboa,, Portugal

Abstract

Background

Metabarcoding is invaluable for understanding trophic interactions, enabling high-resolution and rapid dietary assessments. However, it requires a robust DNA barcode reference library for accurate taxa identification. This dataset has been generated in the framework of the InBIO Barcoding Initiative (IBI) and Agrivole project. The integration of these two projects was crucial, as Agrivole aimed to investigate the trophic niche of small mammals in Trás-os-Montes Region through DNA metabarcoding, which required a reliable plant DNA barcode library for this same region. Given the large number of species not yet represented in international databases, a survey of local plants was essential to fill this gap. Thus, this study created an accurate DNA reference database for the plants of the Trás-os-Montes Region of Portugal.

New information

The current DNA reference database contains 632 vascular plant samples, all morphologically identified and belonging to 435 species. This represents 14% and 38.7% of the total known plant species for Portugal and the study area, respectively.

Of the 1781 barcode sequences provided in this dataset, 1099 contain new information (61.7%) at different levels: 254 (13.6%, ITS2: 41, trnL-ef: 126, trnL-gh: 87) are completely new to GenBank and/or BOLD databases at the time of publication, 438 (24.6%, ITS2: 59, trnL-ef: 173, trnL-gh: 206) are new records for a given species and 407 (22.9%, ITS2: 187, trnL-ef: 206, trnL-gh: 14) provide additional information (e.g. different bp length, intraspecific genetic variability); the remaining 682 sequences (38.3%) are equal (100% identity) to sequences already publicly available for the identified species. Overall, this dataset represents a significant contribution to the genetic knowledge of vascular plants represented in public libraries. This is one of the public releases of the IBI database, which provides genetic and distributional data for several taxa.

All vouchers are deposited in the Herbarium of the Museum of Natural History and Science of the University of Porto (MHNC-UP) and their DNA barcodes are publicly available in the Barcode of Life Data System (BOLD), NCBI GenBank online databases and International Nucleotide Sequence Database Collaboration (INSDC).
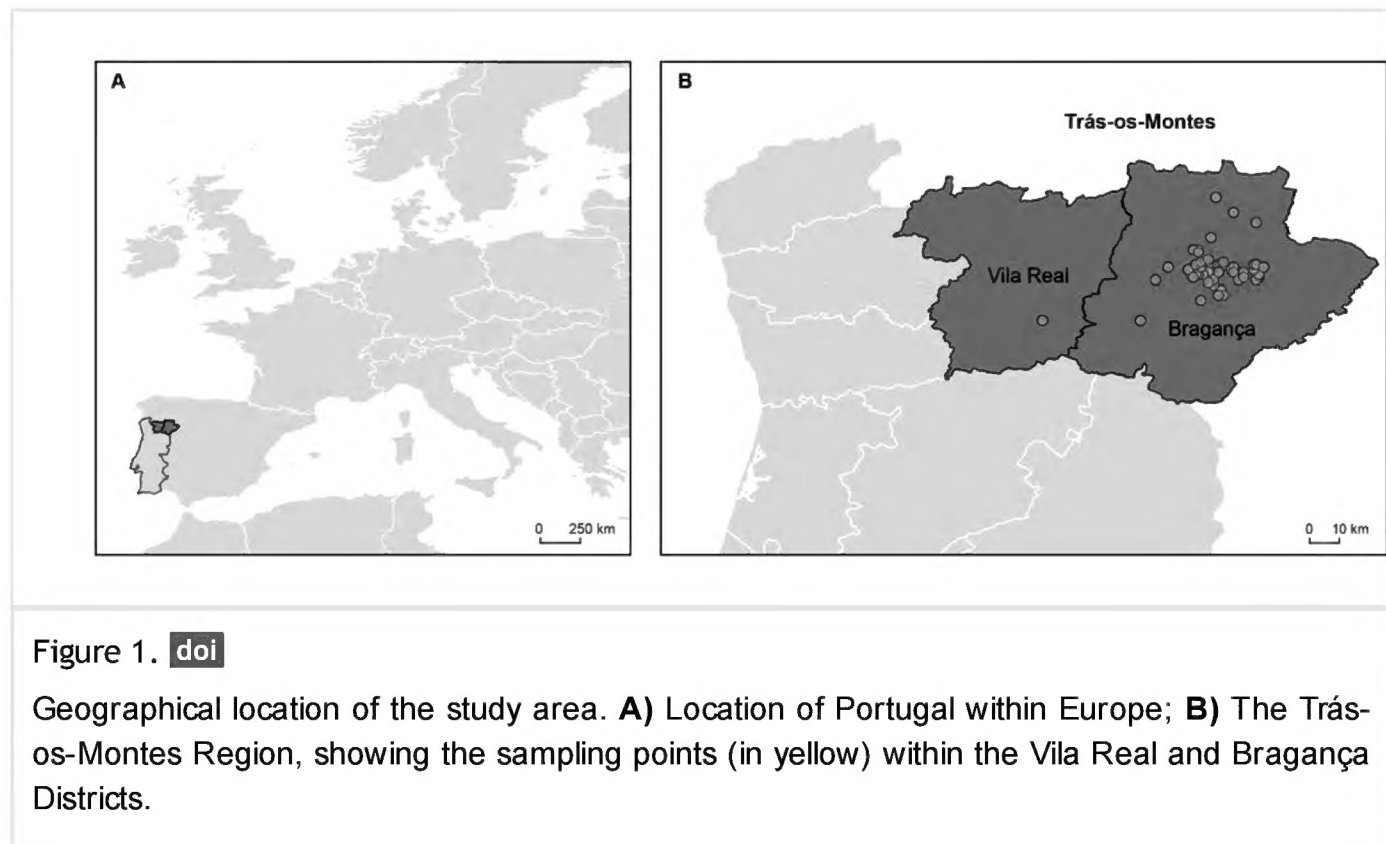
Keywords

Introduction

Traditional approaches to biodiversity assessment have been revolutionised by DNA-based species identification (Hebert and Gregory 2005). DNA barcoding, which uses short standardised regions of DNA, called barcodes, to identify organisms (Hebert et al. 2003), has streamlined the species identification process, making it faster, more automated and less reliant on taxonomic expertise than classical methods (Hebert and Gregory 2005). DNA metabarcoding extends this to complex samples, allowing large-scale taxonomic identification and playing a crucial role in ecological studies, with great applicability to biodiversity conservation (Taberlet et al. 2012, Cristescu 2014, Taberlet et al. 2018). The development of this technique has been facilitated by advances in PCR technologies and high-throughput sequencing (HTS), which allows multiple DNA barcodes to be sequenced from a single sample (Taberlet et al. 2012, Forsman et al. 2022).

Metabarcoding is a particularly valuable tool for understanding trophic interactions (da Silva et al. 2019, de Sousa et al. 2019, Ando et al. 2020), as it enables dietary assessment at a higher taxonomic resolution and is faster than traditional methods and is applicable to non-invasive samples (Thuo et al. 2019, Gil et al. 2020). However, effective identification of taxa from environmental or dietary samples requires a reliable reference library of DNA barcodes (Pompanon et al. 2012, Cuff et al. 2022,Kress et al. 2015, Ferreira et al. 2019). In addition, libraries containing multiple barcode regions can further improve identification by overcoming PCR biases associated with marker selection (da Silva et al. 2019), as overlap between markers increases the likelihood of classifying all taxa in the sample (Fahner et al. 2016). For dietary studies, these accurate reference libraries are crucial for understanding the feeding habits and ecological roles of species (De Barba et al. 2014, Tercel et al. 2021).

The region of Trás-os-Montes is located in the mountainous northeast of Portugal, comprising the Bragança and Vila Real Districts (Fig. 1). This region is characterised by a diverse flora, with numerous endemic species, due to its diversity of habitats and the geochemistry of the substrate (Porto 2020). Two formations in eastern Bragança, for instance, are notable for their basic and ultrabasic igneous rocks, while a carbonate outcrop further east is also of interest (Aguiar and Alves 2020, Aguiar and Monteiro-Henriques 2020). For this reason, several studies have been carried out over the years on the local climate, geology and, especially, on plant diversity and ecology (Sequeira and Silva 1992, Aguiar et al. 2013, Preto et al. 2021). According to *Flora-On: Flora de Portugal Interactiva* (Flora-On: Flora de Portugal Interactiva 2024), of the 3082 plant taxa known for Portugal, 1706 (55.4%) occur in the Trás-os-Montes Region, of which 193 are endemic to the Iberian Peninsula and 22 are endemic to Portugal. Furthermore, Trás-os-Montes harbours 113 threatened species, 15 of which are critically endangered. In order to protect the unique diversity of this region, conservation efforts have been implemented through the establishment of natural parks, including Parque Natural de Montesinho and Parque Natural do Alvão (Queirós 2001). Despite the number of studies focused on this region, the flora of Trás-os-Montes continues to be enriched with new discoveries, of

species previously unknown to occur in Portugal (e.g. Aguiar and Almeida 2006, Araújo 2021, Porto et al. 2024).

Figure 1. `doi`

Geographical location of the study area. **A)** Location of Portugal within Europe; **B)** The Trás-os-Montes Region, showing the sampling points (in yellow) within the Vila Real and Bragança Districts.

The present vascular plant dataset contributes to improving the molecular identification of the flora present not only in Portugal and the Iberian Peninsula, but also in other regions of the world where some species are also present. This DNA barcode reference library was generated in the framework of the InBIO Barcoding Initiative (IBI) and Agrivole project. The IBI project uses HTS technologies to build reference datasets of DNA barcodes from morphologically identified Portuguese specimens (Ferreira et al. 2018). The project has focused mainly on insects due to their importance in food webs and ecosystem functioning, with a total of 26 datasets already publicly avaiable (e.g. Ferreira et al. 2024, Pina et al. 2024). However, the project is now expanding to other groups, including plants. The Agrivole project aims to understand the responses of vole communities in the Trás-os-Montes Region to agroecosystem structure and agricultural practices and how this may affect the potential for pest outbreaks in orchards (i.e. olive groves) or the resilience of vole species of conservation concern in these habitats. The integration of IBI and Agrivole projects was crucial, as the Agrivole project focused on the use of faecal DNA metabarcoding analysis to determine the trophic niche of small mammals from Trás-os-Montes and, in order to obtain accurate and precise dietary estimates, a reliable plant DNA barcoding library for the region was required. As many species are still not represented in open sequence databases, a survey of local plants was essential to fill this gap. Therefore, this study created an accurate DNA reference database for the plants of the Trás-os-Montes Region of Portugal. To achieve this, all plant samples collected were morphologically identified by taxonomic experts and two molecular markers were amplified and sequenced: the chloroplast trnL (UAA) intron region using two primer sets (ef and gh; Taberlet et al. 1991, Taberlet et al. 2007) and the nuclear internal transcribed spacer 2 region, ITS2 using one primer set (UniPlant; Cheng et al. 2016). ITS2 is one of the most widely used molecular markers for plant barcode

identification (Hollingsworth et al. 2011) and trnL is widely used in metabarcoding studies as it can be amplified even when using highly-degraded DNA (Taberlet et al. 2007), making these two regions complementary. This work contributes to the open dissemination and sharing of the distribution records and DNA barcodes of the vascular plant specimens that are part of our reference collection, thereby increasing the available public information. This database will not only serve the Agrivole project, but will also contribute significantly to other DNA metabarcoding studies and traditional molecular research, facilitating ecological, phylogenetic and conservation applications in the region and beyond.

General description

Purpose: This dataset was generated as part of the Agrivole project that aimed to investigate the ecological role of small mammals, particularly voles, in local agroecosystems of the Trás-os-Montes Region in north-eastern Portugal, in collaboration with the InBIO Barcoding Initiative (IBI), which uses HTS technologies to build reference datasets of DNA barcodes from morphologically identified Portuguese specimens. To achieve this goal, the Agrivole project investigated the trophic niche of these small mammals through DNA metabarcoding analysis of their faeces using high throughput sequencing methods. This required a reliable plant DNA barcoding library for the region to accurately and precisely infer the dietary taxa consumed by each species, for which collaboration with IBI was crucial. Therefore, this database aims to provide a reference database of DNA barcodes of the vascular plants of the Trás-os-Montes Region. This library will facilitate the DNA-based identification of plant species in the Agrivole project and other DNA metabarcoding and traditional molecular studies, such as environmental DNA (eDNA) monitoring, which has shown a growing trend in recent years (Cristescu 2014, Fahner et al. 2016, Forsman et al. 2022).

Additional information: A total of 659 vascular plant samples were collected, of which 647 were morphologically identified to the most specific taxon (species, subspecies or variety), comprising 448 species belonging to 69 families (Suppl. materials 1, 2).

Of these, 632 samples were successfully barcoded, covering 435 species (Suppl. material 3). This represents 14% of the known plants in the Portuguese flora and 38.7% of the total known plant species in the study area (considering all the 10 x 10 km UTM squares where the samples are located; Flora-On: Flora de Portugal Interactiva (2024)). The relatively low representativeness is likely due to the high diversity of habitats that exist in the study area and that were not sampled, given the project's focus on the olive groves agroecosytems. To overcome this limitation, future sampling efforts should prioritise the remaining unexplored habitats, such as the basic and ultrabasic igneous rocks or the carbonate outcrops (Aguiar and Alves 2020, Aguiar and Monteiro-Henriques 2020).

With ITS2, 594 samples were successfully barcoded, 590 for trnL primer ef (trnL-ef) and 597 for trnL primer gh (trnL-gh), with 541 samples amplified for the three primer sets. This

resulted in a total of 1781 barcode sequences for the sum of the three markers used (Suppl. materials 1, 4, 5, 6). All 69 families were represented; however, each primerset failed to amplify specific families most likely due to differences in primer specificity, with ITS2 lacking seven families (Aspleniaceae, Dennstaedtiaceae, Dioscoreaceae, Equisetaceae, Ophioglossaceae, Pteridaceae and Cytinaceae), trnL-ef one (Cytinaceae) and trnL-gh three (Montiaceae, Ophioglossaceae and Pteridaceae).

Of the 1781 barcode sequences, 1099 (61.7%) contain new information at different levels: 254 (13.6%, ITS2: 41, trnL-ef: 126, trnL-gh: 87) are completely new to GenBank and/or BOLD databases at the time of publication, 438 (24.6%, ITS2: 59, trnL-ef: 173, trnL-gh: 206) are new records for a given species and 407 (22.9%, ITS2: 187, trnL-ef: 206, trnL-gh: 14) provide additional information (e.g. different bp length, intraspecific genetic variability); the remaining 682 sequences (38.3%) are equal (100% identity) to sequences already publicly available for the identified species. Overall, this dataset represents a significant contribution to the genetic knowledge of vascular plants represented in public libraries and, specifically, within the scope of the Agrivole project, will be essential for studying the feeding ecology of small mammals and understanding their ecological interactions, including competition between species for food resources. This will provide insights into the responses of small mammal communities in the Trás-os-Montes Region to agro-ecosystem structure and agricultural practices and how this may affect the potential for pest outbreaks in orchards (i.e. olive groves) or the resilience of species of conservation concern in these habitats. Furthermore, this database is essential for future long-term monitoring studies of the plant communities in the study area and will assist the government authorities in the sustainable management of the National Parks (Queirós 2001).

Project description

Title: The InBIO Barcoding Initiative Database: contribution to the knowledge of DNA barcodes of the vascular plants of north-eastern Portugal.

Personnel: Joana Paupério (project coordinator), João Queirós (junior researcher of project), Miguel Porto (plant specialist), Paulo Célio Alves (senior researcher in wildlife conservation and management), Pedro Beja (senior researcher in wildlife ecology), Ricardo Pita (project co-coordinator), Hélia Vale-Gonçalves (project technician), Catarina J. Pinho (project technician), Rodrigo Silva (Bachelor student in biology).

Study area description: The region of Trás-os-Montes is located in the mountainous northeast of Portugal, comprising the Districts of Vila Real and Bragança (Fig. 1).

Design description: Vascular plants were collected from traditionally managed olive grove agroecosystems and surrounding areas in north-eastern Portugal. Each distinct vascular plant species within each patch was collected, photographed and then mounted on herbarium sheets *in situ*. These were then morphologically identified by experts and a portion of the leaf tissue was used to obtain DNA barcodes.

Sampling methods

Description: Trás-os-Montes Region, Portugal

Sampling description: The sampling protocol was designed to cover plants in areas where small mammals, particularly voles, were expected to occur, as defined under the Agrivole project (Barão et al. 2022). Plant samples were collected in the northeast of Portugal in August 2018, May–June 2019 and June 2020, from traditionally managed olive grove agroecosystems and their surrounding areas. Olive farming plays a critical role in local economy and has steadily expanded across the region (Duarte et al. 2006), currently covering approximately 15% of the land (unpublished data, Agrivole project). Young olive groves were used as a central point from which a 50 m buffer zone was drawn (Barão et al. 2022). Each distinct vascular plant species within each patch was collected, photographed and then mounted on herbarium sheets *in situ* with an assigned specimen number. A portion of leaf tissue was collected from each sample, placed in a paper envelope and then stored in a zip-lock bag with silica gel until DNA extraction. To minimise contamination between samples, tweezers and scissors were systematically sterilised with 100% ethanol under a flame. After morphological identification, all the specimens (vouchers) were deposited in the Herbarium of the Natural History and Science Museum of the University of Porto (MHNC-UP) with the codes PO-V72113–PO-V72772 (Suppl. material 1)

Quality control: Collected samples were classified and revised by experts in vascular plant identification. The obtained DNA barcode sequences were compared against the GenBank and BOLD databases and the top hits were examined to identify potential PCR / sequencing errors, contaminations or misidentifications.

Step description: 1. Plant sampling

Vascular plants were collected in the northeast of Portugal in August 2018, May–June 2019 and June 2020, from traditionally managed olive grove agroecosystems and their surrounding areas. A leaf tissue sample was collected for DNA extraction.

2. Taxonomic identification

The morphological identification of the specimens was carried out, after herborisation, under a stereomicroscope, using mainly the dichotomous keys of *Flora Iberica* (Castroviejo 2020). The website *Flora-On: Flora de Portugal Interactiva* (Flora-On: Flora de Portugal Interactiva 2024) was used as a reference for the list of Portuguese vascular flora, as well as for visual validation and distribution of the taxa. Regarding nomenclature, the Angiosperm Phylogeny Website (Stevens 2024) was followed for suprageneric taxa, while scientific names were based on the Plants of the World Online (POWO 2024), except where specified otherwise. For some specimens, it was necessary to consult other works namely in cases of possible hybrids (Perez-Carro et al. 1985, Nesom 2006, Portela-Pereira et al. 2013, Krasnopoļska et al. 2020, Király et al. 2021, Martín-Gómez et al. 2022) or, in cases where the *Flora Iberica* has a different treatment of the taxon of

interest than POWO (Pereira Coutinho 1913, Walters 1958, Guittonneau 1972, Alavi 1976 Tutin et al. 1980, Benedí 1986, Snogerup and Snogerup 2001, Uotila 2001, Desfayes 2004, Pavlickf 2007, Horjales et al. 2008, Molina et al. 2008, Blanca 2009, Arabaci et al. 2010, Cano-Maqueda and Talavera 2011, Rubido-Bará et al. 2011, Haerinasab and Rahiminejad 2012, Tison and De Foucault 2014, Särkinen et al. 2018, Karaismailoğlu 2019). Whenever possible, the specimens were identified at the infraspecific level.

3. DNA extraction

DNA was extracted for each sample using 20 mg of dry leaf material weighed into a 2 ml tube. To disrupt the thick plant cell walls, zirconia beads were added to the tube and placed on a mill for a minimum of 10 minutes. Lysis was performed by adding lysis buffer (2% SDS, 2% PVP 40, 250 mM NaCl, 200 mM Tris HCl, 5 mM EDTA, pH 8) and 10 µl of proteinase K (1 mg/ml) to the samples, followed by incubation at 63°C for 30 minutes. RNA was removed by adding 10 µl of RNAase A (10 mg/ml) followed by incubation at 37°C for 20 minutes. DNA was precipitated by adding 150 µl of potassium formate (KCOOH) (3 M), mixed by inversion of tubes, followed by incubation on ice for 25 minutes. A stepwise centrifugation was performed at increasing speed: 1 minute at 1000 rpm, at 2000 rpm, at 4000 rpm, at 8000 rpm and 10 minutes at 11000 rpm, followed by DNA binding using 550 µl of clear supernatant and 825 µl binding buffer (2 M guanidine hydrochloride in 95% ethanol). A maximum of 700 µl of binding mixture was added to a column, followed by centrifugation at 11000 rpm for 1 minute. This procedure was repeated for the remainder of the binding mixture. The column membranes were then washed twice with 80% ethanol (1 minute at 11000 rpm). Finally, two DNA elutions were performed by adding 50 µl of binding buffer (65°C 10 mM Tris-152 HCl, pH 8.3) to the column and centrifuging at 12000 rpm for 1 minute.

4. PCR amplification and library preparation

Polymerase chain reaction (PCR) amplification was performed focusing on two regions, the internal transcribed spacer of nuclear ribosomal DNA (ITS2), using the primer pair UniPlantF/UniPlantR (187–387 bp; Moorhouse-Gann et al. 2018) and the P6 loop of the plastid *trn*L (UAA) gene, which was amplified using two different sets of primers: ef (147–540 bp; Taberlet et al. (1991)) and gh (10–143 bp; Taberlet et al. (2007)). All primers contained Illumina overhang adaptors at the 5' end.

PCR reactions were performed by mixing 5 µl QIAGEN Multiplex PCR Master Mix, 3.4 µl ultrapure water, 0.3 µl of each primer and 1 µl DNA extract. Cycling conditions consisted of an initial denaturation at 95ºC for 15 min, followed by 40 cycles of denaturation at 95ºC for 30 s, annealing at 54ºC for 30 s for ITS2 and 55ºC for 1 min for both *trn*L primers and extension at 72ºC for 30 s, with a final extension cycle at 72ºC for 10 min. A touchdown method was also used for some samples that initially failed amplification for ITS2 to increase the specificity of the PCR product. This was done using the cycling conditions previously presented, changing the annealing for 5 cycles with an initial temperature of 56ºC, reduced by 0.5°C for each cycle. Amplification success was visually verified by electrophoreses in a 2% stained gel agarose using 2 µl of PCR product.

A second PCR was performed to incorporate the Illumina sequencing adapters and P5 and P7 indices using 5 µl KAPA HiFi HotStart ReadyMix, 1 µl of index mix, 2 µl of ultrapure water and 2 µl of the previous PCR products diluted 1:10. Cycling condition consisted of an initial denaturation at 95°C for 3 minutes, followed by 10 cycles at 95°C for 30 seconds, at 55°C for 30 seconds and 72°C for 30 seconds, with a final extension at 72°C for 5 minutes. Indexing success was assessed by comparing the initial PCR product with the indexed product by electrophoresis. These were cleaned using 1.2 × AMPure® XP beads, quantified using Epoch, diluted to 15 nM and pooled by marker. The three libraries were then quantified by qPCR and pooled to obtain a final library at 4 nM. The final library was sequenced on an Illumina MiSeq System, using a MiSeq V2 500-cycle reagent kit and considering a coverage of approximately 5,000 paired-end reads per sample and marker.

5. Bioinformatic analysis

The obtained sequences were bioinformatically processed using the OBITools software (Boyer et al. 2016). Paired-end reads were first aligned using the *illuminapairedend* command and discarded if the overlap quality was less than 40. Unaligned sequences were also removed using the *obigrep* command. Sequences were then dereplicated to unique sequences per sample using *obiuniq* and primer sequences were removed using *ngsfilter*. Sequence lengths were trimmed to the expected lengths mentioned above and sequences with less than two reads were removed. Clustering analysis was performed using *sumacluster* with 99% similarity between sequences to remove PCR and sequencing errors.

6. Data Analysis

To estimate the representativeness of the species collected in this study compared to the expected total diversity known for the Trás-os-Montes Region, we used as the reference dataset the list of plants described by family on the Flora-On website (Flora-On: Flora de Portugal Interactiva 2024).

To assess the contribution of this study as a source of new genetic resources (DNA barcodes), the DNA sequences generated were compared to the NCBI Nucleotide Database using the BLAST+ software (Camacho et al. 2009). The obtained sequences were used to create three local reference databases for each marker (ITS2, trnL-ef, trnL-gh) using the *makeblastDB* function. These databases were then compared to the NCBI online records using the *blastn* algorithm, with a word size of 28 for ITS2 and trnL-ef and a word size of 12 for trnL-gh due to its smaller size, applying minimum cut-offs of 90% identity and 80% coverage. Additionally, the ITS2 sequences were compared against BOLD databases using BOLD Identification Engine (Ratnasingham and Hebert 2007).

Sequences that did not match the morphological identification, at least at the family level, were considered errors and discarded. Sequences that matched with 100% coverage and identity to a different species within the same genus or family were considered new records for that species. Sequences were classified as completely new records for online

databases if they showed identity percentages and query coverage below 90%. Sequences that correctly matched the morphological identification at the species level, but had a coverage of 80%–100% and/or an identity percentage between 98%–100%, were classified as new information records for that species.

Geographic coverage

Description: The study was carried out in the districts of Bragança and Vila Real, located in northeast of Portugal. In particular, vascular plants were collected from agroecosystems.

Coordinates: -7.494 and -6.614 Latitude; 41.832 and 41.336 Longitude.

Taxonomic coverage

Description: A total of 659 specimens were collected, of which 647 were morphologically identified to the most specific taxon (species, subspecies or variety). This dataset includes 448 taxa belonging to 69 families (Suppl. materials 1, 2). Of these, 632 samples were successfully barcoded, covering 435 species from 69 families (Fig. 2; Suppl. material 3). The families Poaceae (15%), Asteraceae (12%) and Fabaceae (12%) represented the highest percentages of samples collected and sequenced, while the remaining 66 families represented less than 5% each.
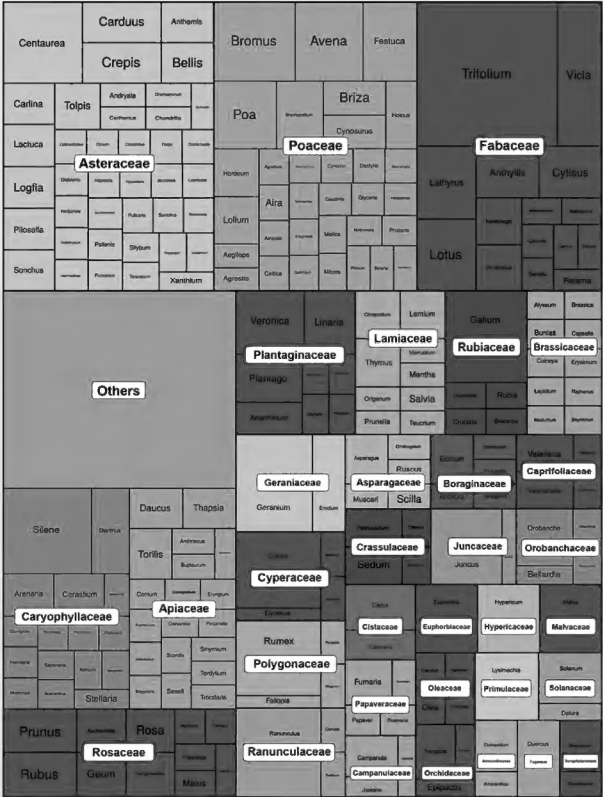

Figure 2. doi

Treemap depicting the genus and corresponding plant families barcoded in this study for the Trás-os-Montes Region. The size of the squares is proportional to the number of species in each genus that were barcoded.

Taxa included:

Rank	Scientific Name	Common Name
kingdom	Plantae	Plants
phylum	Tracheophyta	
class	Liliopsida	
order	Asparagales	
family	Amaryllidaceae	
family	Asparagaceae	
family	Orchidaceae	
order	Dioscoreales	
family	Dioscoreaceae	
order	Poales	
family	Cyperaceae	
family	Juncaceae	
family	Poaceae	
class	Lycopodiopsida	
order	Isoetales	
family	Isoetaceae	
class	Magnoliopsida	
order	Apiales	
family	Apiaceae	
order	Asterales	
family	Asteraceae	
family	Campanulaceae	
order	Boraginales	
family	Boraginaceae	
order	Brassicales	
family	Brassicaceae	
family	Resedaceae	
order	Caryophyllales	
family	Amaranthaceae	

family	Caryophyllaceae
family	Montiaceae
family	Plumbaginaceae
family	Polygonaceae
family	Portulacaceae
order	Cucurbitales
family	Cucurbitaceae
order	Dipsacales
family	Caprifoliaceae
order	Ericales
family	Ericaceae
family	Primulaceae
order	Fabales
family	Fabaceae
family	Polygalaceae
order	Fagales
family	Fagaceae
family	Juglandaceae
order	Gentianales
family	Gentianaceae
family	Rubiaceae
order	Geraniales
family	Geraniaceae
order	Lamiales
family	Lamiaceae
family	Oleaceae
family	Orobanchaceae
family	Plantaginaceae
family	Scrophulariaceae
family	Verbenaceae
order	Malpighiales

family	Euphorbiaceae
family	Hypericaceae
family	Linaceae
family	Salicaceae
family	Violaceae
order	Malvales
family	Cistaceae
family	Cytinaceae
family	Malvaceae
family	Thymelaeaceae
order	Myrtales
family	Lythraceae
family	Onagraceae
order	Piperales
family	Aristolochiaceae
order	Ranunculales
family	Papaveraceae
family	Ranunculaceae
order	Rosales
family	Cannabaceae
family	Rosaceae
family	Urticaceae
order	Santalales
family	Santalaceae
order	Sapindales
family	Anacardiaceae
family	Rutaceae
order	Saxifragales
family	Crassulaceae
family	Paeoniaceae
family	Saxifragaceae

order	Solanales	
family	Convolvulaceae	
family	Solanaceae	
order	Zygophyllales	
family	Zygophyllaceae	
class	Pinopsida	
family	Cupressaceae	
class	Polypodiopsida	
order	Polypodiales	
family	Aspleniaceae	
family	Dennstaedtiaceae	
class	Psilotopsida	
order	Ophioglossales	
family	Ophioglossaceae	
order	Equisetales	
family	Equisetaceae	
family	Pteridaceae	

Temporal coverage

Notes: Plant samples were collected in August 2018, May–June 2019 and June 2020.

Collection data

Collection name: Plantas vasculares - Herbário Português

Collection identifier: Herbário da Universidade Do Porto (PO)

Specimen preservation method: Dry

Curatorial unit: PO-V72113 – PO-V72772

Usage licence

Usage licence: Creative Commons Public Domain Waiver (CC-Zero)

Data resources

Data package title: The InBIO Barcoding Initiative Database: contribution to the knowledge of DNA barcodes of the vascular plants of north-eastern Portugal.

Number of data sets: 1

Data set name: DS-IBPLT

Description: The data underpinning the analysis reported in this paper are deposited at BOLD, the Barcode of Life Data System, with the dataset name DS-IBPLT and GBIF, the Global Biodiversity Information Facility, at: https://doi.org/ 10.15468/exuufa.

Column label	Column description
materialSampleID	Unique identifier for the sample.
recordNumber	Identifier for the sample being sequenced in the IBI catalogue number at Cibio-InBIO, Porto University. Same as fieldNumber.
catalogNumber	Identifier for the specimen deposited in the Museum of the University of Porto (MHNC-UP).
institutionCode	The full name of the institution that has physical possession of the DNA samples.
occurrenceID	Global unique identifier for that sample.
kingdom	Kingdom name.
phylum	Phylum name.
class	Class name.
order	Order name.
subfamily	Subfamily name.
family	Family name.
genus	Genus name.
subpecies	Species nameSubpecies name.
scientificNameAuthorship	The authorship information for the sample.
identificationRemarks	Comments or notes about the taxonomic identification.
identifiedBy	Full name of the individuals who assigned the specimen to a taxonomic group.
typeStatus	Status of the specimen in an accessioning process.
recordedBy	The full names of the individuals or team responsible for collecting the sample in the field.
eventDate	Date of the sample collection.

countryCode	The full, unabbreviated name of the country where the specimens was collectedCode of the country where the specimens was collected.
lat	The geographical latitude (in decimal degrees) of the geographic centre of a location.
lon	The geographical longitude (in decimal degrees) of the geographic centre of a location.
county	The full, unabbreviated name of the county where the organism was collected.
municipality	The full, unabbreviated name of the municipality ("Concelho" in Portugal) where the specimen was collected.
verbatimLocality	The original textual description of the sampled location.
dynamicProperties	DNA barcoded sequences.
measurementDeterminedBy	The full name of the institution where DNA samples are stored.
datasetName	BOLD dataset code.

Acknowledgements

References

- Aguiar C, Almeida JD (2006) *Pimpinella major* (L.) Huds., uma nova Apiaceae para a flora de Portugal. Silva Lusitana 14 (2): 268.
- Aguiar C, Monteiro-Henriques T, Sánchez-Mata D (2013) New contributions on flora and vegetation of northeastern Portugal ultramafic outcrops. Lazaroa 34: 141-150. https://doi.org/10.5209/rev_LAZA.2013.v34.n1.43642
- Aguiar C, Alves P (2020) Afloramentos de carbonatos do Norte de Portugal. In: Porto M (Ed.) Sítios de Interesse Botânico de Portugal Continental. Imprensa Nacional, Lisboa, 229 pp. [ISBN 9722729675].
- Aguiar C, Monteiro-Henriques T (2020) Afloramentos ultramáficos do Nordeste de Portugal. In: Porto M (Ed.) Sítios de Interesse Botânico de Portugal Continental. Imprensa Nacional, Lisboa, 229 pp. [ISBN 9722729675].
- Alavi SA (1976) Genus *Coleostephus* Cassini in Europe (Asteraceae). Phyton, Annales Rei Botanicae 17 (3-4): 319-328.

- Ando H, Mukai H, Komura T, Dewi T, Ando M, Isagi Y (2020) Methodological trends and perspectives of animal dietary studies by noninvasive fecal DNA metabarcoding. Environmental DNA 2 (4): 391-406. https://doi.org/10.1002/edn3.117
- Arabaci T, Dirmenci T, Celep F (2010) Morphological character analysis in Turkish *Micromeria* Benth. (Lamiaceae) species with a numerical taxonomic study. Turkish Journal of Botany 34: 379-389. https://doi.org/10.3906/bot-0905-22
- Araújo PV (2021) Vale do Rio Mente. In: Farminhão J (Ed.) Sítios de Interesse Botânico de Portugal Continental, Tomo II. II. Imprensa Nacional, Lisboa, 285 pp. [ISBN 978-972-27-2967-3].
- Barão I, Queirós J, Vale-Gonçalves H, Paupério J, Pita R (2022) Landscape characteristics affecting small mammal occurrence in heterogeneous olive grove agro-ecosystems. Conservation 2 (1): 51-67. https://doi.org/10.3390/conservation2010005
- Benedí C (1986) Nota sobre *Chamaemelum* Miller. Fontqueria 10: 1-4.
- Blanca G (2009) *Pallenis* Cass. In: Blanca G, Cabezudo B, Cueto M, Fernández López C, Morales Torres C (Eds) Flora Vascular de Andalucía Oriental. 4. Consejería de Medio Ambiente, Junta de Andalucía, Sevilla, Spain, 436 pp. [ISBN 978–84–92807–12–3].
- Boyer F, Mercier C, Bonin A, Le Bras Y, Taberlet P, Coissac E (2016) obitools: a unix-inspired software package for DNA metabarcoding. Molecular Ecology Resources 16 (1). https://doi.org/10.1111/1755-0998.12428
- Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden T (2009) BLAST+: architecture and applications. BMC Bioinformatics 10 (1). https://doi.org/10.1186/1471-2105-10-421
- Cano-Maqueda J, Talavera S (2011) A taxonomic revision of the *Campanula lusitanica* complex (Campanulaceae) in the Western Mediterranean region. Anales del Jardín Botánico de Madrid 68 (1): 15-47. https://doi.org/10.3989/ajbm.2274
- Castroviejo S (2020) Flora iberica: plantas vasculares de la Península Ibérica e Islas Baleares. Real Jardín Botánico, CSIC, Madrid. [ISBN 9788400062224]
- Cheng T, Xu C, Lei L, Li C, Zhang Y, Zhou S (2016) Barcoding the kingdom Plantae: new PCR primers for regions of plants with improved universality and specificity. Molecular Ecology Resources 16 (1): 138-149. https://doi.org/10.1111/1755-0998.12438
- Cristescu M (2014) From barcoding single individuals to metabarcoding biological communities: towards an integrative approach to the study of global biodiversity. Trends in Ecology & Evolution 29 (10): 566-571. https://doi.org/10.1016/j.tree.2014.08.001
- Cuff J, Windsor F, Tercel MTG, Kitson JN, Evans D (2022) Overcoming the pitfalls of merging dietary metabarcoding into ecological networks. Methods in Ecology and Evolution 13 (3). https://doi.org/10.1111/2041-210X.13796
- da Silva L, Mata V, Lopes P, Pereira P, Jarman S, Lopes R, Beja P (2019) Advancing the integration of multi-marker metabarcoding data in dietary analysis of trophic generalists. Molecular Ecology Resources 19 (6). https://doi.org/10.1111/1755-0998.13060
- De Barba M, Miquel C, Boyer F, Mercier C, Rioux D, Coissac E, Taberlet P (2014) DNA metabarcoding multiplexing and validation of data accuracy for diet assessment: application to omnivorous diet. Molecular Ecology Resources 14 (2): 306-323. https://doi.org/10.1111/1755-0998.12188
- Desfayes M (2004) The specific status of *Cyperus badius* and the subspecies of *Scirpoides holoschoenus* (Cyperaceae), with special reference to Sardinia. Flora Mediterranea 14: 173-188.

- de Sousa LL, Silva SM, Xavier R (2019) DNA metabarcoding in diet studies: Unveiling ecological aspects in aquatic and terrestrial ecosystems. Environmental DNA 1 (3): 199-214. https://doi.org/10.1002/edn3.27
- Duarte F, Jones N, Lúcio C, Nunes A (2006) The reform of the olive oil regime and its impacts on the olive and olive oil sector: a case study in Northern Portugal - Trás-os-Montes. New Medit 5 (2): 4-15.
- Fahner N, Shokralla S, Baird D, Hajibabaei M (2016) Large-scale monitoring of plants through environmental DNA metabarcoding of soil: recovery, resolution, and annotation of four DNA markers. PLOS One 11 (6). https://doi.org/10.1371/journal.pone.0157505
- Ferreira S, Fonseca N, Egeter B, Paupério J, Galhardo M, Oxelfelt F, Aresta S, Martins F, Archer J, Corley M, Penado A, Pina S, Jarman S, Beja P (2018) Deliverable 4.2 (D4.2): Protocol for building and organising reference collections of DNA sequences, EnvMetaGen project (Grant Agreement No 668981). Zenodo https://doi.org/10.5281/zenodo.3515967
- Ferreira S, Paupério J, Grosso-Silva JM, Beja P (2019) DNA barcoding of *Sialis* sp. (Megaloptera) in Portugal: the missing tool to species identification. Aquatic Insects 40 (2): 173-184. https://doi.org/10.1080/01650424.2019.1571612
- Ferreira S, Corley M, Nunes J, Rosete J, Vasconcelos S, Mata V, Veríssimo J, Silva T, Sousa P, Andrade R, Grosso-Silva JM, Pinho C, Chaves C, Martins FM, Pinto J, Puppo P, Muñoz-Mérida A, Archer J, Pauperio J, Beja P (2024) The InBIO Barcoding Initiative Database: DNA barcodes of Portuguese moths. Biodiversity Data Journal 12 https://doi.org/10.3897/bdj.12.e117169
- Flora-On: Flora de Portugal Interactiva (2024) Sociedade Portuguesa de Botânica. www.flora-on.pt. Accessed on: 2024-9-30.
- Forsman AM, Savage AE, Hoenig BD, Gaither MR (2022) DNA metabarcoding across disciplines: sequencing our way to greater understanding across scales of biological organization. Integrative and Comparative Biology 62 (2): 191-198. https://doi.org/10.1093/icb/icac090
- Gil V, Pinho C, Aguiar CS, Jardim C, Rebelo R, Vasconcelos R (2020) Questioning the proverb 'more haste, less speed': classic versus metabarcoding approaches for the diet study of a remote island endemic gecko. PeerJ 8: e8084. https://doi.org/10.7717/peerj.8084
- Guittonneau G (1972) Contribuition à l'étude biosystématique du genre *Erodium* L'Her. dans le bassin méditérranéen occidental. Boissiera 20: 1-154.
- Haerinasab M, Rahiminejad M (2012) A taxonomic revision of the genus Trifolium L. sect. Fragifera Koch (Fabaceae) in Iran. Iranian journal of botany 18 (1): 22-30.
- Hebert PN, Cywinska A, Ball S, deWaard J (2003) Biological identifications through DNA barcodes. Proceedings of the Royal Society of London. Series B: Biological Sciences 270 (1512): 313-321. https://doi.org/10.1098/rspb.2002.2218
- Hebert PN, Gregory TR (2005) The promise of DNA barcoding for taxonomy. Systematic Biology 54 (5): 852-859. https://doi.org/10.1080/10635150500354886
- Hollingsworth P, Graham S, Little D (2011) Choosing and Using a Plant DNA Barcode. PLoS ONE 6 (5). https://doi.org/10.1371/journal.pone.0019254
- Horjales M, Laso G, Redondo N (2008) *Dactylis glomerata* L. (Poaceae) en la Península Ibérica. Nova Acta Científica Compostelana (Bioloxía) (17)65-85.
- Karaismailoğlu C (2019) Notes on the taxonomical, morphological and mucilage features of *Capsella bursa-pastoris* Medik. and *Capsella rubella* Reuter taxa (Brassicaceae).

Artvin Çoruh Üniversitesi Orman Fakültesi Dergisi 20 (1): 47-53. https://doi.org/10.17474/artvinofd.475435

- Király G, Hohla M, Nikolić T (2021) Novelties in the vascular flora of Croatia. Natura Croatica: Periodicum Musei Historiae Naturalis Croatici 30 (1): 173-189. https://doi.org/10.20302/NC.2021.30.10
- Krasnopoļska D, Evarts-Bunders P, Svilāne I (2020) Contribution to knowledge of genus *Agrimonia* L. (Rosaceae) in Latvia. Acta Biologica Universitatis Daugavpiliensis 20: 47-53.
- Kress WJ, García-Robledo C, Uriarte M, Erickson D (2015) DNA barcodes for ecology, evolution, and conservation. Trends in Ecology & Evolution 30 (1): 25-35. https://doi.org/10.1016/j.tree.2014.10.008
- Martín-Gómez JJ, Rodríguez-Lorenzo JL, Juan A, Tocino Á, Janousek B, Cervantes E (2022) Seed morphological properties related to taxonomy in *Silene* L. species. Taxonomy 2 (3): 298-323. https://doi.org/10.3390/taxonomy2030024
- Molina A, Acedo C, Llamas F (2008) Taxonomy and new taxa in eurasian Carex (Section Phaestoglochin, Cyperaceae). Systematic Botany 33: 237-250. https://doi.org/10.1600/036364408784571563
- Moorhouse-Gann R, Dunn J, de Vere N, Goder M, Cole N, Hipperson H, Symondson WC (2018) New universal ITS2 primers for high-resolution herbivory analyses using DNA metabarcoding in both tropical and temperate zones. Scientific Reports 8 (1): 8542. https://doi.org/10.1038/s41598-018-26648-2
- Nesom G (2006) *Gamochaeta* (Gnaphalieae). In: Editorial Committee FoNA (Ed.) Flora of North America North of Mexico. 19. Oxford University Press, Oxford, 7 pp.
- Pavlickf LE (2007) *Bromus*. In: Editorial Committee FoNA (Ed.) Flora of North America North of Mexico. 24. Oxford University Press, Oxford.
- Pereira Coutinho AX (1913) *Centaurea*, L. In: Aillaud A, Alves F (Eds) A Flora de Portugal. Livraria Aillaud - Livraria Bertrand, Paris - Lisboa, Rio de Janeiro - S. Paulo - Bello Horizonte.
- Perez-Carro FJ, Fernández-Arece P, Díaz González TE, Salvo AE (1985) Aportación al conocimiento del género *Cheilanthes* en la Península Ibérica. Acta Botanica Malacitana 10: 27-32. https://doi.org/10.24310/Actabotanicaabmabm.v10i.9587
- Pina S, Pauperio J, Barros F, Chaves C, Martins FM, Pinto J, Veríssimo J, Mata V, Beja P, Ferreira S (2024) The InBIO Barcoding Initiative Database: DNA barcodes of Orthoptera from Portugal. Biodiversity Data Journal 12 https://doi.org/10.3897/bdj.12.e118010
- Pompanon F, Deagle B, Symondson WC, Brown D, Jarman S, Taberlet P (2012) Who is eating what: diet assessment using next generation sequencing. Molecular Ecology 21 (8): 1931-1950. https://doi.org/10.1111/j.1365-294X.2011.05403.x
- Portela-Pereira E, Capelo J, Neto C, Costa J (2013) Síntese do Conhecimento Taxonómico do Género *Salix* L. em Portugal Continental. Silva Lusitana 21: 1-55.
- Porto M (2020) Sítios de interesse botânico de Portugal Continental. Imprensa Nacional, Lisboa. [ISBN 9722729675]
- Porto M, Lobo Dias S, Alves P (2024) Three new native species to the Portuguese flora from the rivers of Trás-os-Montes. Acta Botanica Malacitana 49: 128-133. https://doi.org/10.24310/abm.49.2024.17799
- POWO (2024) Plants of the World Online. Facilitated by the Royal Botanic Gardens, Kew. URL: http://www.plantsoftheworldonline.org/

- Preto E, Dias R, Aguiar C (2021) Terras de Trás-os-Montes, uma geologia singular, com rochas raras. Implicações nos georecursos e na flora. Brigantia: revista de cultura 38: 611-620.
- Queirós M (2001) Os parques naturais e o ordenamento do território em Portugal. Inforgeo 16/17: 141-150.
- Ratnasingham S, Hebert PD (2007) bold: The Barcode of Life Data System (http://www.barcodinglife.org). Molecular Ecology Notes 7 (3): 355-364. https://doi.org/10.1111/j.1471-8286.2007.01678.x
- Rubido-Bará M, Horjales M, Villaverde C (2011) Dos nuevas subespecies del género *Jasione* L. (Campanulaceae) en el Noroeste de la Península Ibérica. Nova Acta Científica Compostelana (Bioloxía) 19: 21-31.
- Särkinen T, Poczai P, Barboza G, van der Weerden G, Baden M, Knapp S (2018) A revision of the Old World Black Nightshades (Morelloid clade of *Solanum* L., Solanaceae). PhytoKeys 106: 1-223. https://doi.org/10.3897/phytokeys.106.21991
- Sequeira EM, Silva ARS (1992) Ecology of serpentinized areas of north-east Portugal. In: Roberts BA, Proctor J (Eds) The ecology of areas with serpentinized rocks: a world view. Springer Netherlands, Dordrecht, 28 pp. [ISBN 978-94-011-3722-5]. https://doi.org/10.1007/978-94-011-3722-5_7
- Snogerup S, Snogerup B (2001) *Bupleurum* L. (Umbelliferae) in Europe — 1. The annuals, B. sect. Bupleurum and sect. Aristata. Willdenowia 31 (2): 205-308. https://doi.org/10.3372/wi.31.31201
- Stevens PF (2024) Angiosperm Phylogeny Website. URL: http://www.mobot.org/MOBOT/research/APweb/
- Taberlet P, Gielly L, Pautou G, Bouvet J (1991) Universal primers for amplification of three non-coding regions of chloroplast DNA. Plant Molecular Biology 17 (5): 1105-1109. https://doi.org/10.1007/BF00037152
- Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E (2007) Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. Nucleic Acids Research 35 (3): e14. https://doi.org/10.1093/nar/gkl938
- Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology 21 (8): 2045-2050. https://doi.org/10.1111/j.1365-294X.2012.05470.x
- Taberlet P, Bonin A, Zinger L, Coissac E (2018) Environmental DNA: For biodiversity research and monitoring. Oxford University Press, New York. [ISBN 0198767226] https://doi.org/10.1093/oso/9780198767220.001.0001
- Tercel MTG, Symondson WC, Cuff J (2021) The problem of omnivory: A synthesis on omnivory and DNA metabarcoding. Molecular Ecology 30 (10). https://doi.org/10.1111/mec.15903
- Thuo D, Furlan E, Broekhuis F, Kamau J, Macdonald K, Gleeson D (2019) Food from faeces: Evaluating the efficacy of scat DNA metabarcoding in dietary analyses. PLOS One 14 (12): e0228950. https://doi.org/10.1371/journal.pone.0225805
- Tison JM, De Foucault B (2014) *Allium* L. In: Tison JM, De Foucault B (Eds) Flora Gallica - Flore de France. Biotope Éditions, Mèze, France.
- Tutin TG, Heywood VH, Burges NA, Moore DM, Valentine DH, Walters SM, Webb DA (1980) Flora Europaea. Cambridge University Press [ISBN 0521087171]

- Uotila P (2001) Ranunculaceae. In: Jonsell B, Karlsson T (Eds) Flora Nordica, Volume 2: Chenopodiaceae - Fumariaceae. 2. Royal Swedish Academy of Sciences, Stockholm, Sweden, 1 pp.
- Walters SM (1958) *Montia fontana* L. Watsonia 3: 1-6.

Supplementary materials

Suppl. material 1: IBPLT DNA barcoding of Portuguese plants – Specimen details doi

Authors: João Queirós, Rodrigo Silva, Catarina J. Pinho, Hélia Gonçalves, Ricardo Pita, Paulo C. Alves, Pedro Beja, Joana Paupério, Miguel Porto

Data type: Record information - specimen data

Brief description: The file includes information about all records in BOLD for the IBPLT DNA barcoding of Portuguese plants database. It contains collection and identification data. The data are as downloaded from BOLD, without further processing.

Download file (1.03 MB)

Suppl. material 2: IBPLT DNA barcoding of Portuguese plants – Specimen details – Darwin Core Standard doi

Authors: João Queirós, Rodrigo Silva, Catarina J. Pinho, Hélia Gonçalves, Ricardo Pita, Paulo C. Alves, Pedro Beja, Joana Paupério, Miguel Porto

Data type: Record information - specimen data in Darwin Core Standard format

Brief description: The file includes information about all records in BOLD for the IBPLT DNA barcoding of Portuguese plants. It contains collection and identification data. The data are downloaded from GBIF (**https://doi.org/10.15468/exuufa**), without further processing.

Download file (377.08 kb)

Suppl. material 3: IBPLT DNA barcoding of Portuguese plants – List of specimens and accession codes doi

Authors: João Queirós, Rodrigo Silva, Catarina J. Pinho, Hélia Gonçalves, Ricardo Pita, Paulo C. Alves, Pedro Beja, Joana Paupério, Miguel Porto

Data type: Record information, genetic accession codes

Brief description: List of species DNA barcoded in this project, including their original collection code, BOLD and GenBank accession codes for the three markers used (ITS2, trnl-ef and trnl-gh). In the trnl-gh column, (*) indicates sequences that did not meet the GenBank threshold of > 50 bp sequence length. These sequences are provided in Suppl. material 6.

Download file (52.70 kb)

Suppl. material 4: IBPLT DNA barcoding of Portuguese plants – DNA sequences ITS2 [doi]

Authors: João Queirós, Rodrigo Silva, Catarina J. Pinho, Hélia Gonçalves, Ricardo Pita, Paulo C. Alves, Pedro Beja, Joana Paupério, Miguel Porto
Data type: Genomic data, DNA sequences
Brief description: ITS2 sequences in fasta format. Each sequence is identified by the original collection code, BOLD ProcessID and GenBank/INSDC accession number, separated by a vertical bar.
Download file (192.57 kb)

Suppl. material 5: IBPLT DNA barcoding of Portuguese plants – DNA sequences trnL-ef [doi]

Authors: João Queirós, Rodrigo Silva, Catarina J. Pinho, Hélia Gonçalves, Ricardo Pita, Paulo C. Alves, Pedro Beja, Joana Paupério, Miguel Porto
Data type: Genomic data, DNA sequences
Brief description: trnL-ef sequences in fasta format. Each sequence is identified by the original collection code, BOLD ProcessID and GenBank/INSDC accession number, separated by a vertical bar.
Download file (231.11 kb)

Suppl. material 6: IBPLT DNA barcoding of Portuguese plants – DNA sequences trnL-gh [doi]

Authors: João Queirós, Rodrigo Silva, Catarina J. Pinho, Hélia Gonçalves, Ricardo Pita, Paulo C. Alves, Pedro Beja, Joana Paupério, Miguel Porto
Data type: Genomic data, DNA sequences
Brief description: trnL-gh sequences in fasta format. Each sequence is identified by the original collection code, BOLD ProcessID and GenBank/INSDC accession number, separated by a vertical bar. (*) indicates sequences that did not meet the GenBank threshold of > 50 bp sequence length and therefore do not present accession numbers.
Download file (51.61 kb)